1 Project summary

The casebase framework offers users a convenient method for fitting parametric survival models using generalized linear models (Hanley and Miettinen 2009; Bhatnagar et al. 2022). In this work, an elastic-net multinomial model (enet-Casebase) is proposed for the competing risks scenario (Tamvada 2023). It focuses on examining the performance of the elastic-net penalized multinomial model with casebase sampling and compares it to the elastic-net penalized Cox-Proportional Hazards model. Specifically, variable selection and prediction performance were evaluated in different settings.

The penalized multinomial logistic regression is fitted using stochastic variance reduced gradient descent (SVRG). SVRG shows fast convergence for high-dimensional datasets with a greater number of predictors than observations (p > n) (Johnson and Zhang 2013). This algorithm reduces the variance of stochastic gradients by periodically computing full-batch gradients, referred to as a "snapshot." This gradient adjusts the noisy stochastic gradients used in subsequent iterations.

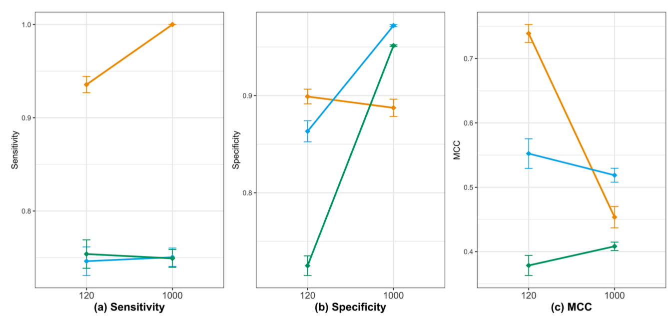
As shown in the Figure 1, the initial results of the proposed method using the SVRG algorithm demonstrate competitive variable selection performance (in terms of Sensitivity, Specificity, and Matthews correlation coefficient) in low-dimensional settings compared to Elastic-net Independent Cox (enet-iCR) and Elastic-net Cox with shared penalty (enet-penCR). In particular, in single-effects settings, enet-casebase exhibits good specificity, sensitivity, and MCC. In both the low and high-dimensional cases, enet-casebase appears to struggle with the opposing effects of covariates. The details of these results can be found in (Johnson and Zhang 2013, 40–45).

Additionally, a two-step procedure was developed to de-bias the estimates of the casebase model and enhance their efficiency. This method demonstrates an ability to accurately predict cumulative incidence comparable to the Cox-Proportional Hazards model and the Fine-Gray model. In the Figure 2, it is evident that the de-biased casebase fit yields estimates that are smooth over time.

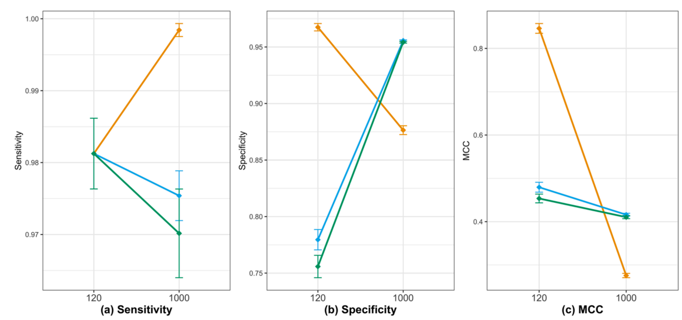
From the previous results, it can be noted that the elastic net multinomial model does not perform well on high-dimensional data. After reviewing the optimization process, it can be observed that SVRG achieves suboptimal coefficients in those settings. Therefore, we have implemented an Accelerating Variance-Reduced Stochastic Gradient algorithm to achieve better optimization in approximately the same amount of time. This accelerated algorithm combines SVRG's variance reduction with momentum to speed up convergence (Driggs, Ehrhardt, and Schönlieb (2022)). Like SVRG, this algorithm starts with a full-batch gradient computation. However, instead of relying solely on stochastic updates, it introduces a sequence of extrapolated points, which are constructed using a weighted combination of the current and previous iterations. As a result, incorporating this algorithm enables obtaining more optimal coefficients in a similar amount of time compared to using SVRG.

1.1 A brief example

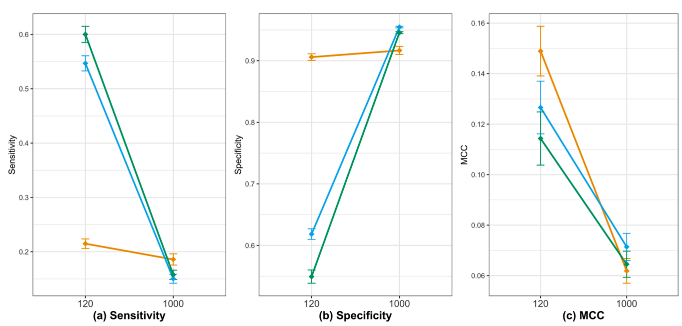
This section analyzes results from the elastic-net multinomial model for competing risks using various optimization algorithms. Data are simulated from a $K = 2$ competing risks proportional hazards model. The cause-specific hazard for cause for individual i with covariates $X_i = (X_{i1}, \ldots, X_{ip})$ at time t is $\lambda_k(t|X_i) = \lambda_{0k}(t)\exp(X_i^T \beta_k)$. The baseline hazards $\lambda_{0k}(t)$ follow a Weibull distribution $\lambda_{0k}(t) = h_k \gamma_k t^{\gamma_k - 1}$, with parameters $(h_1, \gamma_1) = (0.55, 1.5)$ and $(h_2, \gamma_2) = (0.35, 1.5)$.



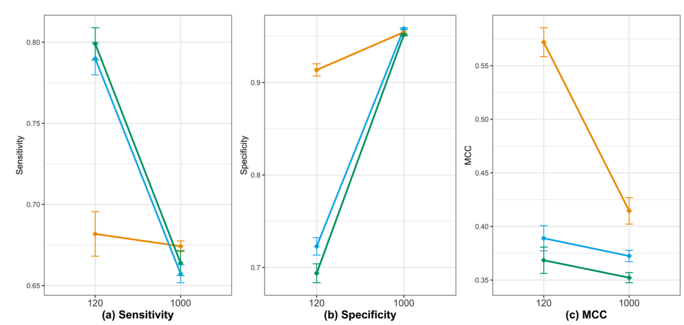
(a) Single effects on endpoint of interest



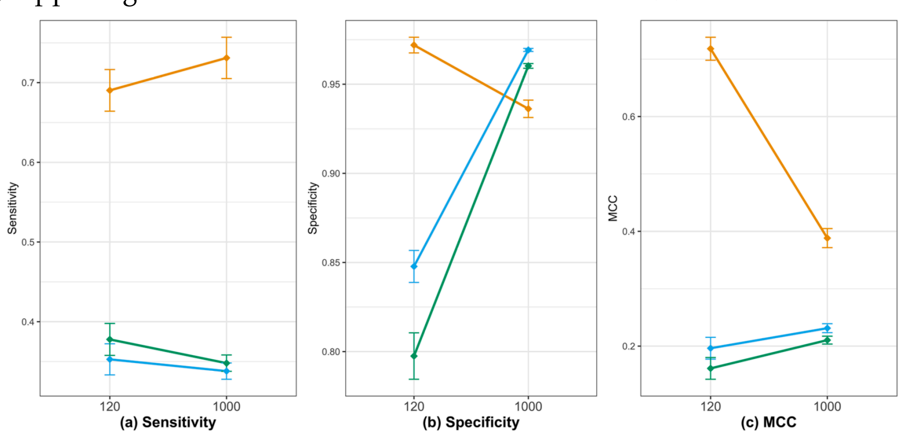
(b) Single effects on both endpoints



(c) Opposing Effects

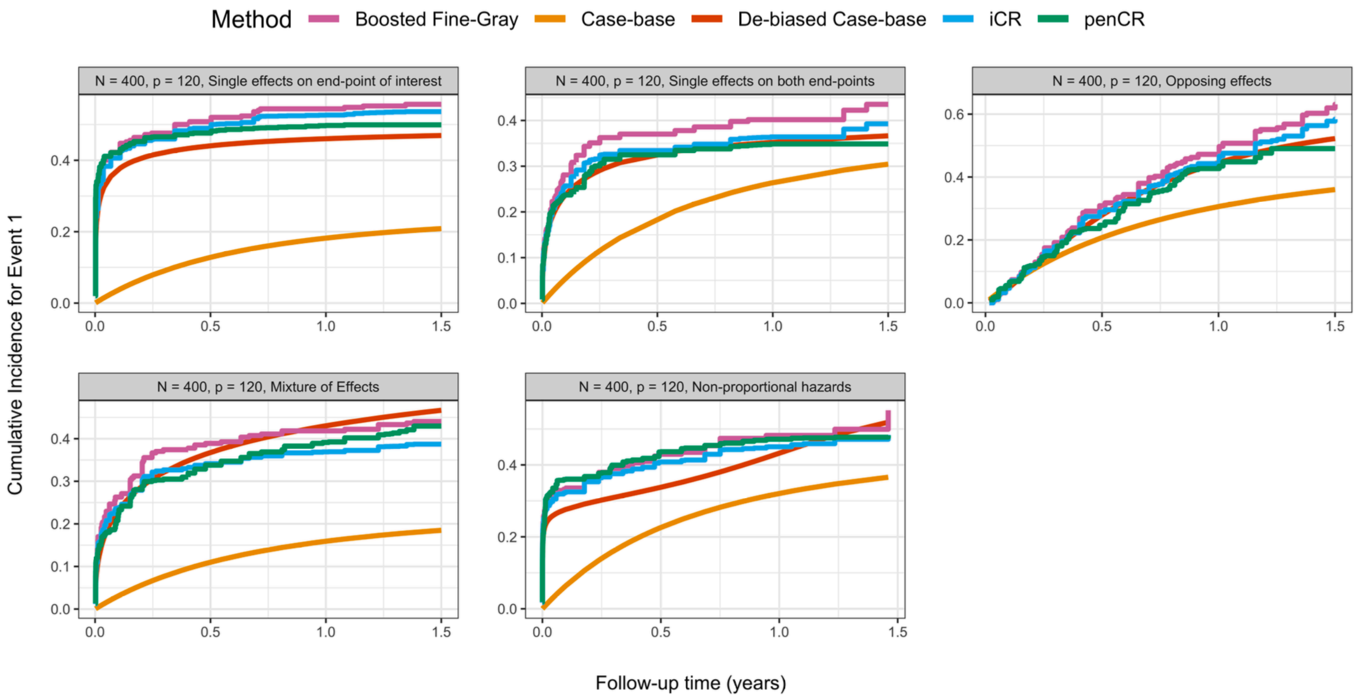


(d) Mixture of effects

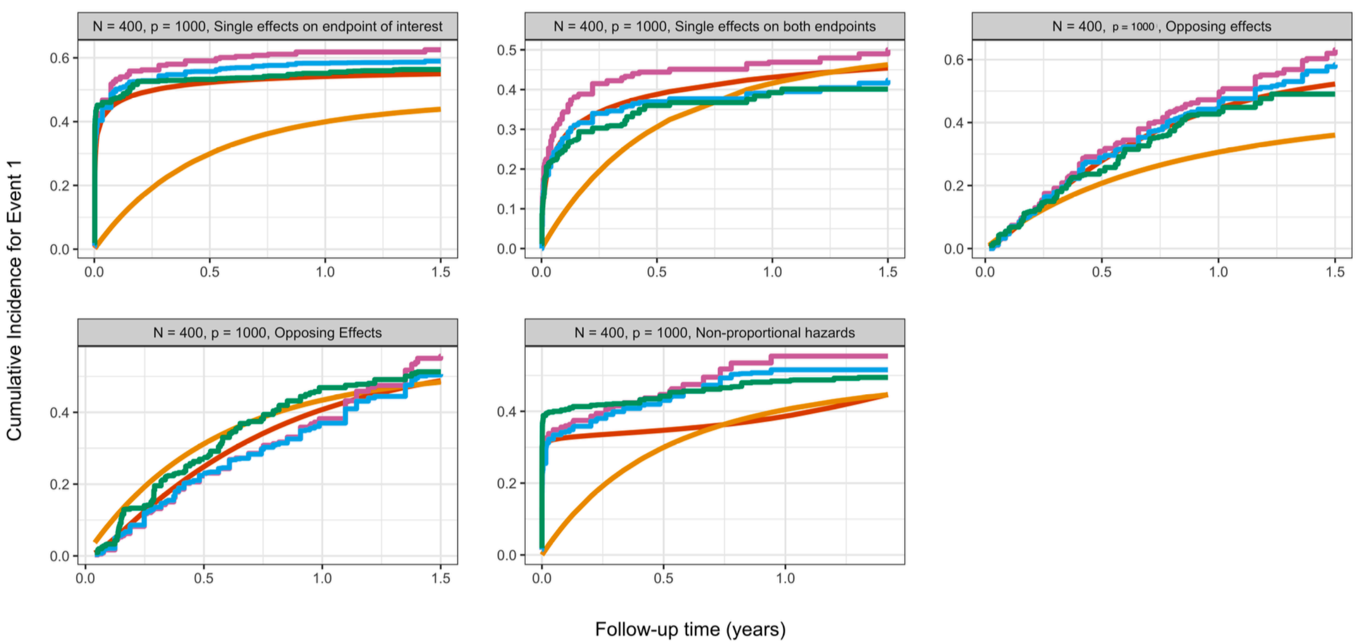


(e) Non-proportional cause-specific hazards

Figure 1: Findings from Initial Experiments



(a) Low-dimensional case ($n = 400$, $p = 120$)



(b) High-dimensional case ($n = 400$, $p = 1000$)

Figure 2: Cumulative incidence curves for event 1

The coefficient vectors $\beta_1, \beta_2 \in \mathbb{R}^p$ are sparse, with non-zero effects predominantly within the first 18 predictors. For cause 1, the coefficients for $X_1, \ldots, X_{18}$ are set as $(1, 1, 1, 1, 1, 1, 0.5, -0.5, 0.5, -0.5, 0.5, -0.5\ 1, 1, 1, 1, 1, 1)$, and 0 for $j > 18$. For cause 2 (β_2), the coefficients for $X_1, \ldots, X_{24}$ are set as $(0, 0, 0, 0, 0, 0, 0.5, -0.5, 0.5, -0.5, 0.5, -0.5, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1)$, and 0 for $j > 24$.

Event times T_i and causes C_i are generated by simulating potential failure times T_{ik} from $\lambda_k(t|X_i)$ and setting $T_i = \min(T_{i1}, T_{i2})$ with C_i being the index k for which $T_{ik} = T_i$. Independent censoring times $T_{cens,i}$ are generated based on an overall rate of 0.05. Observed data consist of $(ftime_i, fstatus_i)$, where $ftime_i = \min(T_i, T_{cens,i})$ and the status $fstatus_i = C_i \cdot \mathbb{1}(T_i \leq T_{cens,i})$ (with $fstatus_i = 0$ indicating censoring).

Given the discussed setting, 15 simulations were generated for this example. For each simulation, the elastic-net multinomial model was estimated using the SVRG algorithm with the default tolerance, a version with smaller tolerance, and the Accelerated SVRG algorithm. The penalization parameter λ in all cases was tuned using 10-fold cross-validation. Figure 3 reports the selection performance for each case based on the λ that minimizes the deviance. The elasticnet penalty parameter α is set to 0.7 to promote sparsity. As a result, it can be observed that the accelerated version tends to produce sparser models and achieve higher MCC scores, particularly for the cause of interest.

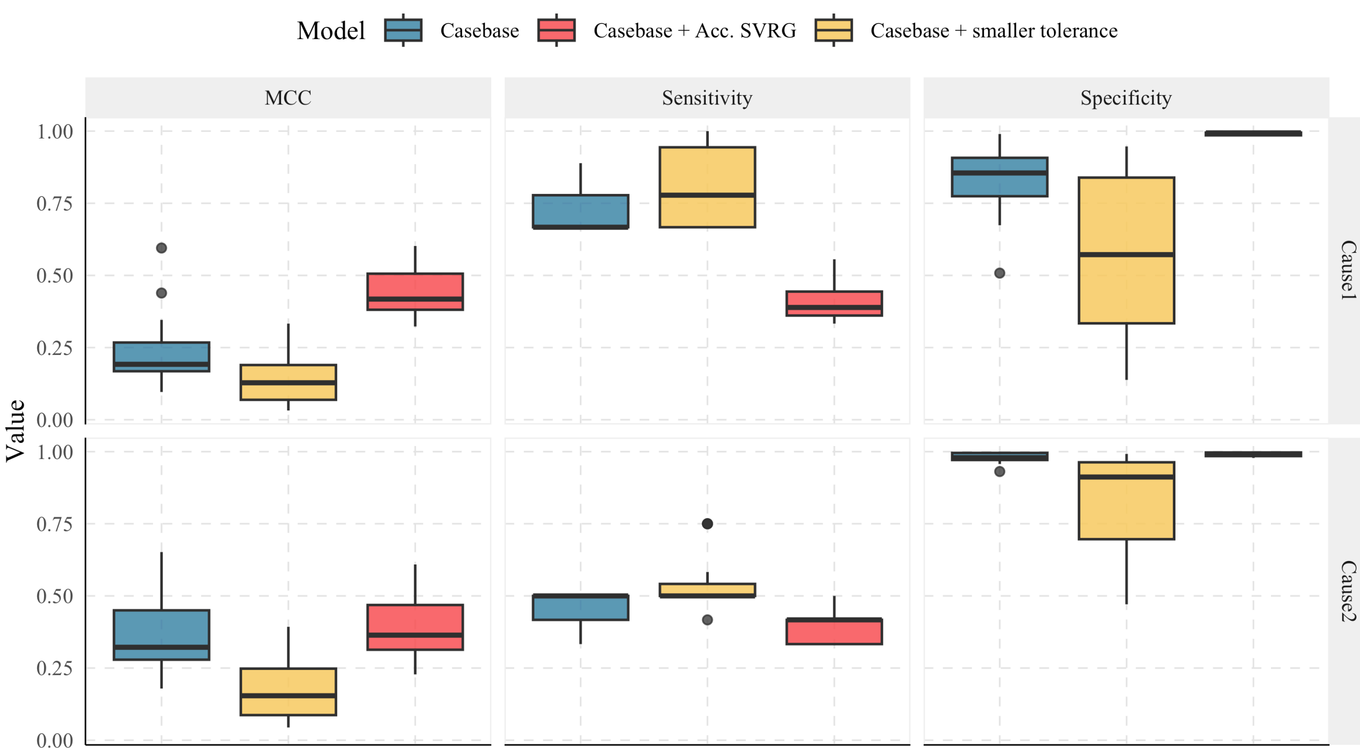


Figure 3: Sensitivity, Specificity and MCC Across Models

Now, defining the Mean Squared Error (MSE) as $\frac{1}{p}\sum_{k=1}^{p}(\hat{\beta}_k - \beta_k)^2$, the Figure 4 presents the results derived from this formula across 15 simulations for each model. As observed, the estimated parameters tend to be closer to the true parameters, indicating improved accuracy.

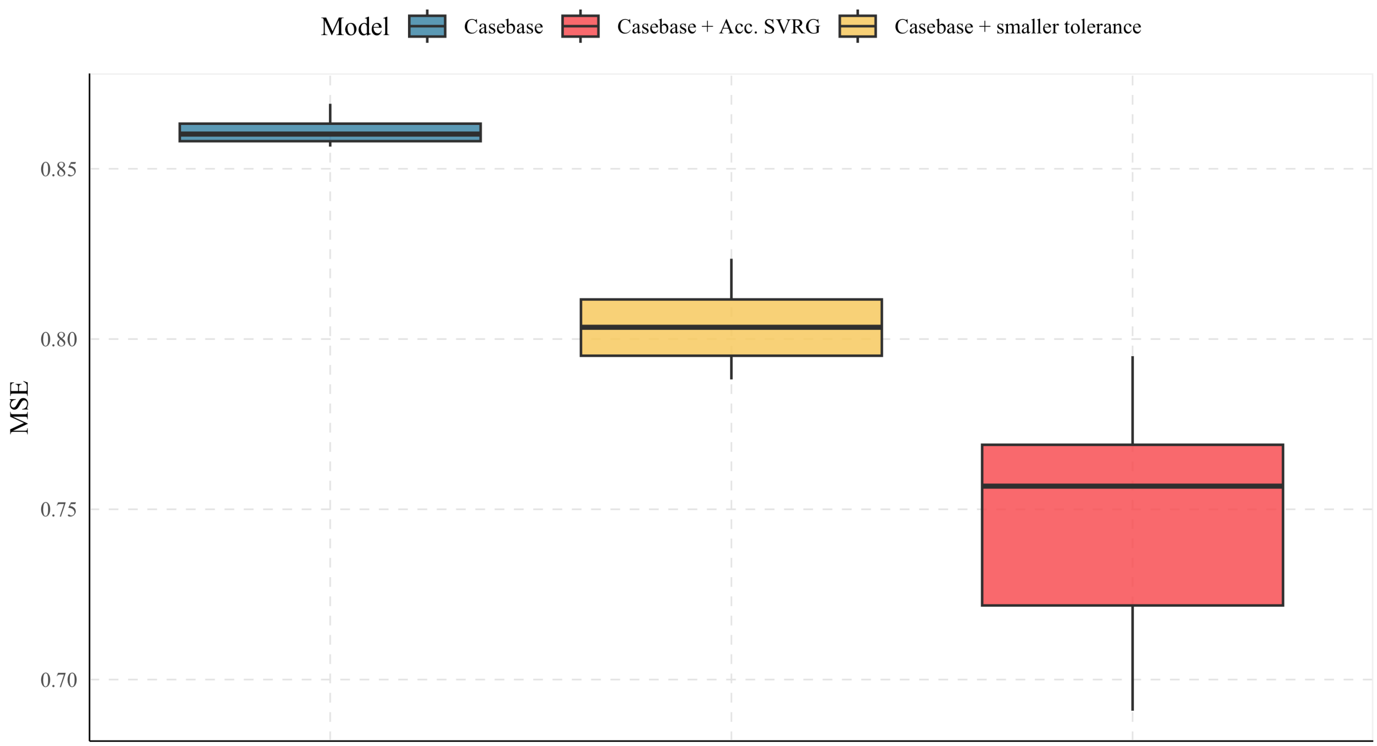


Figure 4: MSE Across Models

1.2 Next Steps

Since the accelerated optimization of the elasticnet multinomial model demonstrated an improvement in high-dimensional settings, the next steps focus on rerunning the experiments presented in Tamvada (2023). Afterwards, the model will be evaluated using a real-world dataset for competing risk.

References

Bhatnagar, Sahir Rai, Maxime Turgeon, Jesse Islam, James A. Hanley, and Olli Saarela. 2022. "Casebase: An Alternative Framework for Survival Analysis and Comparison of Event Rates." *The R Journal* 14 (3): 59–79. https://doi.org/10.32614/RJ-2022-052.

Driggs, Derek, Matthias J. Ehrhardt, and Carola-Bibiane Schönlieb. 2022. "Accelerating Variance-Reduced Stochastic Gradient Methods." *Mathematical Programming* 191 (2): 671–715. https://doi.org/10.1007/s10107-020-01566-2.

Hanley, James A., and Olli S. Miettinen. 2009. "Fitting Smooth-in-Time Prognostic Risk Functions via Logistic Regression." *The International Journal of Biostatistics* 5 (1). https://doi.org/10.2202/1557-4679.1125.

Johnson, Rie, and Tong Zhang. 2013. "Accelerating Stochastic Gradient Descent Using Predictive Variance Reduction." In *Advances in Neural Information Processing Systems*. Vol. 26. Curran Associates, Inc. https://papers.nips.cc/paper_files/paper/2013/hash/ac1dd209cbcc5e5d1c6e28598e8cbbe8-Abstract.html.

Tamvada, Nirupama. 2023. "Penalized Competing Risks Analysis Using Casebase Sampling." University of British Columbia. https://doi.org/10.14288/1.0435526.